DOCUMENT RESUME

ED 157 910 TM 007 214

AUTHOR Wagner, Michael; Rohwer, William D., Jr.
TITLE Premise and Inference Memory as a Function of Age and Context.
PUB DATE [Mar 78]
NOTE 27p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Age Differences; *Cloze Procedure; Concept Formation; *Context Clues; Elementary Secondary Education; Grade 5; Grade 11; Learning Theories; *Logical Thinking; *Memory; Performance Factors; *Recognition; Semantics

ABSTRACT

A sentence completion task was used to investigate age differences in childrens' ability or inclination to invoke premises and inferences during prose processing, tasks which (according to the constructive hypothesis) adults typically perform. A pilot study confirmed that the sentence completion task was preferable to the recognition paradigm, because the latter does not differentiate between subjects who remember premises, inferences, or integrated representations of an array. Participants, 112 subjects each from grades five and eight, were tested for premise and inference memory. The study design varied the following between-subject factors: (1) information (premise, inference)--to independently assess retention; (2) grade (fifth, eighth)--to verify age differences; (3) conditions--to vary degree of contextual support for inference making; and (4) test lists (eight lists)--for counterbalancing. On inference, as expected, eleventh graders scored higher than fifth graders. This pattern of performance was not altered with the addition of context. No age differences were evident for the premise task; both age groups were highly proficient. Contrary to the constructive hypothesis, the benefits of supplying context were not limited to the younger group. Furthermore, neither age group performed above chance expectations on inferences in the standard condition. (Author/CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Premise and Inference Memory as a Function of Age and Context[1]

Michael Wagner

and

William D. Rohwer, Jr.
University of California, Berkeley

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Michael Wagner

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ED157910

TM007 214

Abstract

A sentence completion task was employed to assess memory for premises and inferences as a function of age and context. The sentence completion task was preferable to the recognition paradigm for interpretive reasons. On inferences, as expected, eleventh graders scored higher than fifth graders. This pattern of performance was not altered with the addition of context. No age differences in performance were evident for the premise task, in which both age groups were highly proficient. These findings are evaluated with respect to predictions which can be derived from the constructive hypothesis.

[1]This research was supported in part by grants from the National Institutes of Health (HD03869) and N.I.G.M.S. (GM 01207-14). The report was prepared at the Institute of Human Learning, which is supported by grants from the National Institutes of Health.

We are grateful to Bill Graff, to the principals, Hal Ruby and Cynthia Scortt, to the teachers and students of the Orinda Union School District, and to the students, teachers, and principal, Bran Yaich, of Miramonte High School, for their participation and assistance in conducting this research.

Premise and Inference Memory as a Function of Age and Context

Bransford and others (Bransford and Franks, 1971; Barclay, 1973; Bransford and Franks, 1973) have advanced the idea that adults typically use the information in prose materials to construct semantic descriptions of situations. Moreover, experimental evidence (Bransford, Barclay, and Franks, 1972; Paris and Mahoney, 1974; Wagner and Rohwer, 1977) has supported the view that this integrative process can occur across sentence boundaries and results in the psychological representation of between sentence or inference information.

The question remains whether this constructive hypothesis must be modified to accommodate age differences in the ability or inclination to invoke such integrative processes. While several studies have supported the notion that children integrate semantic information as well as adults (Paris and Carter, 1973; Paris and Mahoney (verbatim condition), 1974; Brown, 1976; Paris, 1975), conflicting evidence exists as well (Paris and Mahoney (paraphrase condition), 1974; Paris and Upton, 1976; Hildyard and Olson, 1977; Wagner and Rohwer, 1977).

Most of the studies cited above share a common methodology, that is, the recognition paradigm. Results produced by this paradigm, however, have been susceptible to alternative interpretations for one or more of the following reasons: the false recognition paradigm uses errors, which typically decrease with age, as an index of inference making (Lindauer and Paris, 1976); verbatim test assertions which preserve premise noun order

require only lexical information to score high on the inference index; and, when the within-story premise order is held constant (AB-BC), application of a simple rule can result in high inference performance.

More importantly, serious questions can be raised concerning the efficacy of using any form of the recognition paradigm to explore inference making in a three-term series problem. The input in such a task typically consists of pairs of comparative or locative relational premises (e.g., The coin weighed more than the ring. The ring would sink faster than the nail.). After input presentation subjects are asked to classify a set of premise (e.g., The ring was lighter than the coin.) and inference (e.g., The coin would sink more quickly than the nail.) paraphrased test assertions as being consistent or inconsistent with the input.

In this situation, suppose a recognition test list, composed entirely of inference items, were administered. Further, suppose the subjects can remember only a single premise relationship from each story in the input and thus are incapable of deriving any inferences from them. If the subjects adopt a random guessing strategy their performance should not differ from the chance expectations. However, their guessing strategy may not be entirely random. Given the story presented above, suppose a subject remembered only the first premise. When presented with the inference test assertion, "The coin would sink faster than the nail," the subject might reason that, since it is consistent with the premise he remembers, insofar as in both sentences the coin weighed more than the entity to which it was compared, a yes response would be considered correct. Indeed, a single premise from any story would be sufficient for a subject using this strategy to respond correctly to any inference test assertion, regardless of either the order or position

of the premises in the input, or the order or validity of the inference test assertions. By using this strategy consistently, therefore, a subject could perform as well or better on inferences than on premises.

Subtle methodological variations across studies may systematically influence the propensity of non-inference makers to invoke such a strategy, resulting in varying effects. Thus, there appear to be inherent ambiguities in the recognition measures which may preclude the independent assessment of memory for inference information.

The assessment of memory for premise information with a recognition paradigm may also suffer from interpretive ambiguity. By definition, memory for a premise consists of memory for the specific elements and the relationship that comprise a single proposition or sentence. Thus, a subject who remembers the premises in a three-term series task should be able to distinguish between valid and invalid premise test assertions. In addition, he should be able to distinguish between a valid premise and a valid inference because the inference represents a new combination of elements, albeit ones consistent with the premises. Furthermore, if semantic differences exist in the relations which connect the premise pairings of entities, and the subject remembers both premises from the input, it seems that he should be able to distinguish between them on the basis of the relations used to compare the entities (e.g., weighs more than, could sink faster than).

The inadequacy of the recognition paradigm with regard to the measurement of the retention of premise information becomes evident when it is realized that no discrimination among target premises, alternative premises, and inferences is required of subjects in deriving an index of memory for premises. In fact, a subject, acting only on the basis of memory for inferences, could

correctly classify premises by acting consistently with the inferences. Thus, if a subject who remembered that "The coin would sink faster than the nail" was presented with the premise, "The coin weighed more than the ring," he could respond so as to be consistent with the redundant information he possesses, that is, that the coin weighed more than whatever it was compared to. In fact, he would always be correct. Consequently, a correct response to a premise assertion does not insure that the subject remembers that specific premise.

If a subject represented the input as a single integrated array (e.g., coin>ring>nail), he should be able to distinguish quite easily between premises and inferences. However, without remembering information specific to each premise, he may be unable to distinguish between the premises.

In conclusion, the premise recognition task, which simply requires validity discriminations, may not differentiate between those subjects who remember premises, inferences, or integrated representations of the array.

It is clear, then, that a new task must be devised which overcomes the interpretive difficulties presented by the recognition paradigm. One task which appears promising is the sentence completion task.

The sentence completion task has the apparent advantage over the recognition task in the measurement of memory for inferences insofar as the subject, rather than being forced to make a decision regarding an inference, has the opportunity to respond confidently according to whatever kind of information he has available in memory. An inference test sentence for the example given earlier might be "The _____ would sink more quickly than the other two things." If the subject failed to make an inference and thus responded on the basis of premise information, he might confidently

respond with "ring" since the second premise in the input asserted that the ring would in fact sink faster than the nail. Alternatively, he might respond "coin" since the first premise asserted that the coin weighed more than the ring. Of course, for a subject acting primarily on the basis of premise information, the response "ring" might be more appealing since the relation used in the test sentence is more congruent with the relation used in the second premise than with the relation used in the first premise of the input. Of course, if the subject made the correct inference, which takes into account both premises, he would answer "coin." The advantage, then, of this task is that memory for a single premise from any story cannot guarantee that the proportion of correct responses will exceed the response rate to the inappropriate premise ("ring") choices.

The sentence completion task may also offer advantages over the recognition paradigm in the assessment of memory for premises. While the recognition task only requires subjects to make validity judgments, the sentence completion task requires, in addition, that subjects discriminate target premises from both alternative valid premises and valid inferences.

In the premise sentence completion task a test sentence for the example given earlier might be "The _____ would sink more quickly than the _____." The subject is informed that the test assertion is a paraphrase of one of the input sentences and that he should fill in the blanks in the test sentences so as to preserve the meaning of that target sentence. The response can then be classified as valid or invalid, as a premise or an inference, and, if a premise, as the target or the alternative premise. Thus, in order to correctly respond, the subject, in addition to overall meaning of the story, must possess information specific to the target sentence (<u>was heavier</u> versus <u>would sink faster than</u>). The advantage of this task is

that an wholistic representation of the array or memory for an inference by itself cannot insure that the production rate of valid target responses will exceed that of alternative valid premises or valid inferences.

The principal issue being addressed in this study is whether age differences exist in the ability or inclination to invoke integrative activities during text processing. One account of such age differences stems from the hypothesis that age differences in inference making emerge when the input provides insufficient support to prompt the younger subjects to represent the key premises within a framework or context. In order to evaluate this conception, two versions of each input passage can be generated: a baseline version (Standard) and a Context-Added version. In the Context-Added version, several sentences can be added to the baseline version in order to supply an actor, a beginning situation, an action, and a suggested resolution to the situation, thus serving to provide a context for the key relational premises. According to the hypothesis, then, benefit from such a prompted version of the input should be most pronounced for the younger sample.

A second issue addressed in this study regards the consequences of semantic integration for the retention of premise information. According to the constructive hypothesis (Barclay, 1973), premise information is subsumed by wholistic descriptions of situations, and, thus, one premise should be indiscriminable from another consistent one (presented or not) or from an inference (presented or not). Early support was garnered for this view by Bransford and Franks (1971) but recently their conclusions have been disputed (Flagg, Potts, and Reynolds, 1975; Flagg, 1976; Reynolds and Flagg, 1976; Lawson, 1977). The disagreement on this issue may arise from the

fact that indexes of memory for premise information in studies of semantic integration have simply required subjects either to discriminate valid from invalid assertions, whether premise or inference, or to apply a verbatim criterion in a situation made ambiguous by the unequal number of old and new instances in the test list. Thus, the indices used in the above studies appear to be inconsistent with the kind of measures required to assess the constructive hypothesis. This gap may be filled by the sentence completion task.

In order to test the efficacy of the sentence completion method, a pilot study was conducted. The design included the between-subjects factors of Grade (fifth, eleventh), Context (standard, context-added), Response task (recognition, sentence completion) and, for the sentence completion task, Information (premise, inference). In addition, the recognition task included the within-subjects factors of Information (premise, inference) and Validity (valid, invalid).

Eight paragraphs, each comprised of six sentences, were composed so that the second and fifth sentences were the key relational premises. The first key premise provided a relation between two entities (A-B), while the fifth sentence related the third entity with one of the first two (B-C). Each premise could be presented in two orders (A-B or B-A, B-C or C-B), resulting in the generation of four alternative orderings for the key premises, each represented by two paragraphs in the input list. The Context-Added stories were generated by adding four sentences (CA) to the Standard versions (S) in the following manner: CA, S-1, CA, S-2, S-3, S-4, CA, S-5, S-6, CA. Forty-eight fifth and forty-eight eleventh grade students from a single upper-middle class suburban population participated in the study.

The dependent measures for the recognition test were the proportion of correct responses to valid and invalid test assertions. For the inference portion of the sentence completion task the dependent variable was the proportion of correct responses and the number of "inappropriate premise" responses. For the premise portion of the task, the production rates of valid target premises, alternative premises, and inferences served as the dependent variables.

While the recognition task failed to produce any significant main effects or interactions, some interesting patterns were visible. In the Standard condition, there were no age differences in premise performance, but eleventh graders scored higher on inferences than did fifth graders. The age difference on the inference measures was reduced in the Context-Added condition as a result of improved performance by the fifth graders. A similar pattern of inference results was found in the sentence completion task, but the effects were of much greater magnitude, consistent with the notion that the recognition paradigm lacks sensitivity to age differences in inference making.

In the premise portion of the sentence completion task, all groups gave target premises as responses (overall mean proportion = .46) far more often than either alternative premises (.15) or inferences (.16). This finding appears to be incompatible with the constructive hypothesis (Barclay, 1973).

The findings of this pilot study support the use of the sentence completion method and indicate that age differences exist in the propensity to make inferences from prose materials when minimal context is provided. It also appears that, when such inference making occurs, contrary to pre-

dictions derived from the constructive position, retention of information specific to single sentences is not precluded. In order to confirm and extend these pilot findings, a formal study was designed, using the sentence completion method, with the purpose of assessing the constructive hypothesis developmentally.

In particular, on the inference task it was expected that, for the input condition with the greatest contextual support, both older and younger subjects would produce more inference responses than would be expected on the basis of chance. Furthermore, for input conditions with less contextual support, it was expected that the performance of the younger subjects would suffer relative to the older subjects.

This study was also designed to evaluate the effects of constructive processing on the retention of premises. According to the constructive position, as integrative activities increase, whether as a result of increasing age or increasing contextual support in the input, the capability of distinguishing target premises from alternative premises and inferences should diminish.

Method

Subjects

All participants in this study were drawn from two elementary schools and one high school serving the same upper-middle class suburban residential area. Fifth graders were randomly selected from two elementary schools. Eleventh graders were randomly selected from those volunteers enrolled in required U.S. history classes. One hundred and twelve subjects from each grade level comprised a total sample of 224.

Materials

Eight core paragraphs, each comprised of six sentences, were constructed. An example is displayed in Table 1. These paragraphs, which served as the input for the Standard condition, were composed so that the second and fifth sentences were always the key relational premises from which an inference could be derived.

The first sentence in each paragraph introduced the three entities (A, B, C) to be compared. The order in which the entities appeared in this sentence was random. The second sentence stated a relationship on a comparative dimension (e.g., weighed more than) between two of the entities (A relation-1 B). The fifth sentence extended this relationship by comparing one of the entities mentioned in the first sentence with the third entity (B relation-2 C). The result is that an inference could be derived regarding the relative positions of the two entities that were not directly compared (A, C).

In order to prevent subjects from using lexical information in making the inference, the relation in the latter premise was never a verbatim repetition of the relation in the former premise. In fact, the relations in the two premises were never even direct paraphrases (e.g., weaker, had less strength) of one another. While the relation in the first premise (e.g., John was weaker than Bill) compared two entities directly on the underlying comparative dimension (e.g., strength), the relation in the second premise must be inferred from the underlying dimension (e.g., Bill could lift less weight than Mike). The other three sentences comprising a paragraph serve as fillers. They serve to either identify attributes of the entities (The ring was made of copper) or to locate the entities (The ring was in the box).

The key relational premises were constructed with the intent that they would be equally and highly plausible so that, in order to correctly respond to a sentence completion question, the meaning of the study sentences would have to be represented in memory.

Four comparative dimensions were included in the eight paragraphs (weight, size, strength, wealth). Each dimension was represented twice in all study lists. For each paragraph a marked and unmarked form (Clark, 1969) were generated. These were counterbalanced across lists such that within any single list each dimension was represented by one paragraph in marked form and another in unmarked form.

Furthermore, if the entities in the key relational premises were always presented in the same order (i.e., A-B, B-C), the subjects might discover a simple rule which would allow them to make the correct inference. In order to prevent this from occurring, four presentation orders of the entities were derived (A-B, B-C; B-A, B-C; A-B, C-B; B-A, C-B). Consequently, eight lists were required to counterbalance for markedness and premise order.

The Standard-Added form of the input was derived by simply adding three filler sentences to the Standard version as shown in Table 1.

The Context-Added version was derived by altering the initial sentence in the Standard version and by adding three sentences in the same positions as those added to create the Standard-Added form of the input. These alterations were carried out with the aim of adding thematic elements to the Standard form of the paragraph. Specifically, an initial situation involving the three entities was described in the first sentence. This situation implied some action relevant to the dimension on which the entities were compared. The other three sentences, rather than serving as

fillers, were intended to lead the listener towards some resolution of the initial situation (i.e., identifying the object that would sink the fastest).

Finally, the Context-Cued input was generated by simply adding a sentence prior to the Standard paragraph. This sentence explicitly directed the listener to attend to the relevant dimension in the paragraph and attempted to involve the listener in a process (e.g., find out how fast they sink) that would lead to the inference (e.g., the coin will sink faster than the ring or the nail).

The relational terms used in the input paragraphs were never repeated in the corresponding test sentences, whether premise or inference, to prevent the use of lexical information in identifying correct responses. An inference test assertion was generated by either paraphrasing or using the antonym of the relation in the second key premise of a paragraph (B relation-2 C) so that the correct response would always be the entity (A) mentioned only in the first key premise. For example, an inference test sentence for the paragraph in Table 1 would be derived by paraphrasing relation-2 (i.e., The _____ would sink more quickly than the other two things.). Thus, a subject acting on the basis of a single premise, rather than taking both key premises into account, might make a response consistent with the second premise (e.g., The ring would sink faster than the nail) but not with the first (e.g., The coin was heavier than the ring.).

Premise test sentences always included the same relationships as those appearing in corresponding inference test sentences (e.g., The _____ would sink more quickly than the _____.). The relational terms in half of the test sentences were paraphrases of the relation-2s in the input while the remaining relational terms were antonyms of the relation-2s.

Both premise and inference test lists contained eight test sentences, one representing each paragraph in the input. The order in which the paragraphs were presented at study was random and was the same for all subjects. The order in which paragraphs were tested was a different random order and again was the same for all subjects.

Procedure

The participants were tested individually in a mobile laboratory. Assignment to experimental conditions was random with the exception that half as many subjects were assigned to the Context-Cued condition as to each of the other three conditions. The task consisted of an acquisition and test phase. During acquisition subjects were read a practice paragraph. Following this, subjects were told to listen carefully to the set of eight paragraphs, played back to them on a tape recorder, and that they would later be tested on what they remembered. The paragraphs had been prerecorded on cassette tape in a male voice at the rate of two words per second followed by a blank interval such that the total presentation time for each paragraph was 35 seconds. The paragraphs, which were numbered, were played on a portable tape recorder.

Following the acquisition phase, subjects were asked a practice question to insure their adoption of the appropriate criterion, premise or inference. They were given feedback on the questions, regardless of the correctness of their responses. Administration of test instructions took about two minutes. The subjects were then given a typed copy on a single page of the appropriate premise or inference test list. The three entities in a paragraph always accompanied the test sentence about that paragraph. The experimenter read the questions to the subject and noted the subject's oral responses.

Design

The design of the study can be summarized by citing the four between-subjects factors. Information (premise, inference) was included to permit independent assessment of the retention of premise and inference information. Secondly, Grade (fifth, eleventh) was included for the purpose of verifying the existence of age effects in inference making. The third factor, Conditions (Standard, Standard-Added, Context-Added, Context-Cued), was intended to vary the degree of contextual support for inference making supplied in the input. Finally, the factor of Test lists (eight lists) was included for counterbalancing purposes.

Results

On the inference task, the variables used as an overall index of performance were the proportion of inference and "inappropriate premise" responses. For the premise task, three variables were used to assess performance: proportions of valid target premise, valid alternative premise, and valid inference responses. The results are displayed in terms of these variables in Table 2.

Inference Task

In order to determine whether performance exceeded chance expectations on the inference task, an analysis of variance was used to compare the values of the inference task variables with 0.33. The analysis revealed that inference responses were selected at a higher rate than would be expected on the basis of chance ($F(1,48) = 27.70$, $p < .05$) while inappropriate premise responses were not ($F = 3.27$). This advantage in the rate of choosing inference over inappropriate premise responses is consistent with the

constructive hypothesis. Indeed, if subjects were responding primarily on the basis of premise information, no advantage could be expected in favor of inference responses.

It remains to be seen, however, whether age differences exist in the ability or inclination to invoke such integrative processing activities. In order to evaluate this possibility, the analysis of variance was extended to include the between-subjects factors of Grade and Conditions. The main effect of the first factor, Grade, was significant for inference responses ($F(1,48) = 4.23$, $p < .05$) but not for inappropriate premise responses ($F < 1$). These findings are consistent with the expectation of age differences in the application of constructive processes.

In an attempt to isolate the locus of this Grade effect, inference performance was compared with chance within individual cells. For fifth graders the only condition in which the proportion of inference responses significantly exceeded 0.33 was the Context-Cued condition ($F(1,48) = 8.59$, $p < .05$). For eleventh graders, on the other hand, inference responses exceeded chance in both the Standard-Added ($F(1,48) = 12.77$, $p < .05$) and Context-Cued ($F(1,48) = 18.51$, $p < .05$) conditions. While neither fifth nor eleventh graders appear to have made inferences in the Standard condition, increasing contextual support appears to produce more inference making in both groups ($F(1,48) = 10.58$, $p < .05$). Thus, in the Context-Cued condition both age groups performed well above chance expectations. Despite the appearance of a Grade difference in the pattern of inference performance across conditions, the Grade X Conditions interaction was not significant ($F < 1$).

Premise Task

In order to determine whether performance exceeded chance expectations

on the premise task, an analysis of variance was used to compare the values of the premise task variables with 0.17. The response rate for valid target premises exceeded chance expectations ($F(1,48) = 222.99$, $p < .05$) as did that for valid alternative premises ($F(1,48) = 11.19$, $p < .05$) while the valid inference rate did not ($F < 1$).

A comparison of the two grade levels did not detect significant effects for target premises ($F = 1.29$), alternative premises ($F = 2.03$), or inferences ($F = 3.17$). Evidently the advantage of eleventh graders over fifth graders on the inference task does not hold on the premise task.

Similarly, the improvement in inference performance with increasing contextual support was not general to premise performance. None of the comparisons among conditions were significant for either the alternative premise or inference variables. In fact, for target premises, contextual support significantly impeded performance, as shown by the comparison of the Standard with the Context-Added conditions ($F(1,48) = 4.58$, $p < .05$). The contrast between the Standard and Context-Cued condition was in the same direction but failed to reach significance ($F = 1.51$). Despite this trend, subjects in all conditions at both Grade levels produced more valid target premises than would be expected on the basis of chance.

Derived Variables

According to the constructive position, an increase in integrative activities should be accompanied by a corresponding decrease in the ability to distinguish valid target premises from either valid alternative premises or valid inferences (Barclay, 1973). In order to test this hypothesis, two variables are derived. The first variable, inference discrimination, was generated by taking the difference between the mean production rate of valid

premises (average of targets plus alternatives) and inferences. Thus, if a subject failed to remember information pertaining to the particular combination of entities comprising premises in the input, inferences should be produced as often as either of the premises. Consequently, the inference discrimination variable would not differ from zero. The second variable, premise discrimination, was derived by taking the difference between the production rate of valid target and valid alternative premises. This premise discrimination variable should reflect the subject's memory for semantic information specific to the target premise. Thus, subjects who store wholistic representations of the input should be unable to distinguish between the valid premises. In this situation, the value of the premise discrimination variable should not differ from zero. The results are presented in terms of these variables in Table 3.

An analysis of variance carried out on these derived variables indicated, contrary to predictions drawn from the constructive hypothesis, that both variables differed significantly from zero, and that this difference held for all eight groups of subjects defined by the factors of Grade and Conditions. Thus, both fifth and eleventh graders could distinguish between premises and inferences and, further, could distinguish between the target and alternative premises.

In defense of the constructive position it could be argued that the inference task was quite difficult for both age groups and, to a large extent, precluded inference making. However, according to this position, the inference advantage that eleventh graders had over fifth graders should be reflected in a relative disadvantage in premise performance, especially as indexed by the derived variables. No such Grade differences were found

for either the inference discrimination variable ($F < 1$) or for the premise discrimination variable ($F = 2.12$). Surprisingly, there was a trend on the premise discrimination variable in favor of the eleventh graders in the opposite direction of that expected. While the eleventh graders were superior to fifth graders at answering inference questions correctly, there was no evidence of a corresponding deficit in the eleventh graders' ability to distinguish target premises from alternative premises or inferences.

Although this finding is difficult to reconcile with the constructive position, it is possible that failure to find the predicted deficit on the derived premise measures for the eleventh graders may represent a general performance factor related to age. In order to evaluate this conception a more sensitive test of the constructive position may be conducted with reference to changes in the value of the derived premise variables relative to the amount of contextual support in the input conditions. Thus, since inference performance for both Grades improved with increasing contextual support, a corresponding reduction in performance on the derived premise measures would be consistent with the constructive position. This prediction, while not supported for the inference discrimination variable ($F's < 1$), was confirmed for the premise discrimination variable (Standard versus Context-Added, ($F(1,48) = 4.87$, $p < .05$)) consistent with the constructive position.

Despite this confirming evidence, the finding that all subjects, regardless of Grade or Condition, were able to distinguish valid target premises from valid alternative premises and valid inferences is incompatible with the constructive position. In order to resolve this issue, input conditions must be devised which lead participants to higher performance levels on the inference task, allowing a stronger test.

Discussion

The pattern of results obtained in this study on the inference task differed in two respects from the predicted pattern of results. First of all, it was expected that the effects of supplying Context for the input would benefit younger subjects more than older subjects. While the findings indicated that there was a significant benefit of enhancing Context, the benefit was not limited to the younger group.

Secondly, it was predicted that older subjects would perform at a higher level than younger subjects on the Inference task, and that this difference should be greatest in the Standard condition. In other words, the older subjects were considered to be less dependent on the provision of Context for integrative processing to be initiated. However, contrary to predictions, neither age group performed above chance expectations on inferences in the Standard condition. In fact, the pattern of results was such that, with the addition of Context, older subjects appeared to take greater advantage of integrative activities than younger subjects, although this outcome was not confirmed.

These findings may be understood if the difficulty of the Standard form of the input was such that it precluded even the older group from applying constructive processes. In fact, the error rate for the inference task did not differ from chance expectations for all subjects in the Standard condition ($F = 1.89$). Thus, contrary to our initial predictions, the addition of context could not be expected to benefit the younger more than the older subjects.

Finally, the present results bear on the question how integrative activities affect the retention of premise information. The prediction

was that the retention of semantic information specific to the target premise would decrease as a function of increasing age and increasing contextual support to the point that valid target premises would be indistinguishable from either valid alternative premises or valid inferences. Partial confirmation was obtained from the contextual manipulation, but the results from the age manipulation were disconfirming. In addition, in no instance were valid target premises indistinguishable from other valid assertions. However, inference performance did not reach a level which would have afforded a satisfactory test of these hypotheses. Yet, despite such interpretive problems, the sentence completion test paradigm appears to represent a satisfactory approach for the assessment of the constructive hypothesis developmentally and for determining the fate of premise information when constructive processes are invoked.

References

Barclay, J.R. The role of comprehension in remembering sentences. Cognitive Psychology, 1973, 4, 229-254.

Bransford, J.D., Barclay, J.R., & Franks, J.J. Sentence memory: A constructive vs. interpretive approach. Cognitive Psychology, 1972, 3, 193-209.

Bransford, J.D. & Franks, J.J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

Bransford, J.D. & Franks, J.J. The abstraction of linguistic ideas: A review. Cognition: International Journal of Cognitive Psychology, 1973, 211-249.

Brown, A.L. Semantic integration in children's reconstruction of narrative sequences. Cognitive Psychology, 1976, 8, 247-262.

Clark, H.H. Linguistic processes in deductive reasoning. Psychological Review, 1969, 76, 387-404.

Flagg, P.W. Semantic integration in sentence memory? Journal of Verbal Learning and Verbal Behavior, 1976, 15, 491-504.

Flagg, P.W., Potts, G.R., & Reynolds, A.G. Instructions and response strategies in recognition memory for sentences. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 592-598.

Hildyard, A. & Olson, D.R. Children's abilities to draw inferences from oral material. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

Lawson, R. Representation of individual sentences and holistic ideas. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 1-9.

Lindauer, B.K. & Paris, S.G. Problems with a false recognition paradigm for developmental memory research. Journal of Experimental Psychology: Human Learning and Memory, 1976, 22, 319-330.

Paris, S.G. Integration and inference in children's comprehension and memory. In F. Restle et al. (Eds.), Cognitive Theory, Vol. 1. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc., 1975

Paris, S.G. & Carter, A.Y. Semantic and constructive aspects of sentence memory in children. Developmental Psychology, 1973, 9. 109-113.

Paris, S.G. & Mahoney, G.J. Cognitive integration in children's memory for sentences and pictures. Child Development, 1974, 45, 633-642.

Paris, S.G. & Upton, L.R. Children's memory for inferential relationships in prose. Child Development, 1976, 47, 660-668.

Reynolds, A.G. & Flagg, P.W. Recognition memory for elements of sentences. Memory and Cognition, 1976, 4, 422-432.

Wagner, M. & Rohwer, W.D., Jr. Developmental changes in prose memory. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

Table 1: Example of a Standard Paragraph and its Three Alternative Forms (Standard-Added, Context-Added, Context-Cued)

Standard	The coin, nail, and ring were all shiny. The coin weighed more than the ring. The ring was made of copper. The coin was a quarter. The ring would sink faster than the nail. The nail was smooth.
Standard-Added	The coin, nail, and ring were all shiny. The ring was in a box. The coin weighed more than the ring. The ring was made of copper. The coin was a quarter. It was in a purse. The ring would sink faster than the nail. The nail was smooth. It was under a board.
Context-Added	John was testing how fast a coin, nail, and ring would sink. He threw the ring in the pond. The coin weighed more than the ring. The ring was made of copper. The coin was a quarter. John also threw the coin in the pond. The ring would sink faster than the nail. The nail was smooth. John picked it up.
Context-Cued	For this story imagine that you were throwing some objects in a pond to find out how fast they sink.

Table 2: Mean Proportion Correct for Inference and Premise Tasks as a Function of Grade and Conditions.

		Infer. Task				Premise Task			
Grade	Condition	Infer-ence	Inapp. Prem.	Target Prem.	Alter. Prem.	Infer-ence	Inval. Target	Inval. Alter.	Inval. Infer.
Fifth	Standard	.37	.37	.50**	.20	.05	.07	.13	.05
	St. Add.	.35	.35	.41**	.22	.09	.11	.12	.05
	Con. Add.	.37	.38	.37**	.26**	.08	.12	.08	.09
	Con. Cue.	.50*	.39	.41**	.20	.09	.20	.08	.02
Elev.	Standard	.39	.33	.49**	.18	.14	.06	.10	.03
	St. Add.	.47*	.35	.51**	.15	.09	.14	.06	.05
	Con. Add.	.40	.41	.41**	.22	.09	.08	.12	.08
	Con. Cue.	.58*	.28	.44**	.23	.11	.11	.08	.03

*Mean proportion significantly above chance, or 0.33 (α = .05)

**Mean proportion significantly above chance, or 0.17 (α = .05)

Table 3: Mean Values of the Derived Premise Variables as a Function of Grade and Conditions

Grade	Condition	Premise Discrimination	Inference Discrimination
Fifth	Standard	.30*	.30*
	St. Add.	.19*	.23*
	Con. Add.	.11*	.23*
	Con. Cue.	.20*	.25*
Elev.	Standard	.31*	.19*
	St. Add.	.36*	.23*
	Con. Add.	.19*	.23*
	Con. Cue.	.20*	.23*

*Mean value significantly above chance, or 0.00 ($\alpha = .05$)